

EXHIBIT B

Contributions of Human Tumor Xenografts to Anticancer Drug Development

Edward A. Sausville and Angelika M. Burger

Marlene and Stewart Greenebaum Cancer Center, and Department of Pharmacology and Experimental Therapeutics, University of Maryland School of Medicine, Baltimore, Maryland

Abstract

Mouse models of cancer have consistently been used to qualify new anticancer drugs for study in human clinical trials. The most used models include transplantable murine tumors grown in syngeneic hosts and xenografts of human tumors grown in immunodeficient mice. For the latter systems, retrospective preclinical-clinical correlation studies are available, which suggest that improvements must be made to increase their value. Transgenic, knock-out, and knock-in mouse models and their intercrosses are more recent developments that mirror defined steps of human carcinogenesis. However, their value in predicting clinical results remains to date poorly defined. We take the position that properly used and interpreted human tumor xenografts grown in immunodeficient mice can be useful, although not absolutely predictive of behavior in the clinic, and continue to make contributions to critical clinical development choices. (Cancer Res 2006; 66(7): 3351-4)

Background

A controversy exists in defining the optimal animal model for assessing activity of a proposed drug for development as an anticancer agent. The basis for this dispute lies in an insufficient predictive power of the traditionally explored tumor model systems for how actual human beings will respond to the treatment in the clinic (1, 2). This is in contrast to other disease areas (e.g., fungal infections), where evidence of a protective preclinical effect in rabbits readily translates into clinically active agents, assuming that human pharmacology and toxicology can be optimized (3).

Even when drugs with evidence of anticancer activity in preclinical *in vivo* models are given at their maximum tolerated doses, they frequently fail to produce useful activity in humans. At one extreme, "animal model nihilists" take the position that any animal model of cancer is intrinsically flawed, and some actually advocate proceeding to human trials based on hypothetical or *in vitro* supporting data. At the other extreme, advocates of genetic models of cancer tout transgenic models, sometimes constructed with multiple genetic lesions, as desirable in providing a comparable model to the human disease (1, 2, 4, 5).

Key Findings

Historically, cancer drug screening efforts relied for several decades on mouse tumors, particularly leukemia systems, such as L1210 or P388, to define active agents. The U.S. National Cancer

Institute (NCI) promoted such an effort from the mid 1950s, adding mouse solid tumor models as they became available (Colon 38, B16 melanoma, Lewis lung, and M5076 reticulosarcoma) in the 1970s through the beginning of the 1980s (5-7). A frequent concern, however, was that the murine tumor models are intrinsically flawed in their value to predict activity in human disease. This is exemplified by the fact that these models had identified 35 therapeutic agents until the early 1980s when it became evident that the classes of anticancer drugs found active against mouse tumors comprised mainly alkylating and other DNA-damaging agents, which are also toxic to the bone marrow (our "classic" arsenal of cytotoxics) and that novel structures had not been discovered for >20 years (6, 7). Reasons for the deficiency of syngeneic mouse models are the limited variety of available tumor types and a rapid growth, with average doubling times of <2 days (7). In addition, biological agents to be studied fairly in such models would need to be produced in a species-directed way. Finally, there are examples of compound action that seem to have intrinsically different features in mouse tumors or cells as opposed to their human cognates: compound classes, such as brefeldins (8) or certain minor groove DNA binders (9), seem to have lesser activity in murine tumors or murine normal cell compartments than in those derived from humans. Nonetheless, transplantable mouse tumors remain valuable today because of an intact tumor host environment that allows for the evaluation of therapies that require immune response or that target-specific components of blood vessels or the extracellular matrix (5, 7). An example from this category of agents is sunitinib, a complementarity-determining region-grafted humanized antibody directed against human fibroblast activation protein, which is a cell surface antigen of reactive tumor stromal fibroblasts (10). Sunitinib was pre-clinically developed based on the murine homologue owing to the lack of appropriate human model systems (10).

The availability of athymic mice (*nu/nu*) and subsequent immunodeficient mouse strains with other genetic lesions (severe combined immunodeficiency, SCID) allowed by the mid-1980s the widespread possibility of testing human tumor explants and cell lines grown as xenotransplants for response to cancer drugs (refs. 6, 7, 11; Fig. 1). Such models raised the hope that they would reveal agents more likely to have activity in solid human cancers. One of their strengths is the broad spectrum of available tumor types (patient explant and cell line-derived models have been described for all major histologies) and the possibility of *ex vivo* genetic or therapeutic manipulation before xenotransplantation (refs. 6, 7, 11; Fig. 1). However, a number of artificial features associated with their use must be acknowledged: the blood supply and neovascularization is provided by the host; the stroma is murine; orthotopic transplantation is technically difficult; therefore, one mostly selects for tumors occurring in an artificial tissue compartment, the subcutaneous site. As result of s.c. transplantation, metastatic disease (i.e., that aspect of the carcinogenic

Requests for reprints: Edward A. Sausville, Marlene and Stewart Greenebaum Cancer Center, University of Maryland School of Medicine, Room 59D07, 22 South Greene Street, Baltimore, MD 21201. Phone: 410-328-7394; Fax: 410-328-6896; E-mail: esausville@umm.edu

©2006 American Association for Cancer Research.
doi:10.1158/0808-5472.CAN-05-3627

With this backdrop, how could such models possibly be considered valuable? One immediate issue that must be addressed is where in the "natural history" of a drug's development is the animal model being used? If anticancer activity in the animal model is to be pursued before pharmacology of the agent is optimized, or with vehicles, such as DMSO (owing to nonoptimized pharmaceutical features), which impede systemic administration, then it is clear that a basis for false reporting of a compound's prospects exists. In two recent positive retrospective reviews (14, 15) of xenograft models, either agents were used where the pharmacology in humans is well understood, and therefore, doses and schedules can be selected that realistically mirror human use in the case of pediatric neoplasms (15) or human tumors were transplanted directly from patients into the mouse (14) and not passaged as cell lines. The theoretical advantages of the latter approach is that the xenograft histology closely resembles that of the patient's tumor (i.e., implying that the tumor cells rather than the host stroma dictate the tumor architecture, molecular characteristics are maintained, and a spectrum of proangiogenic features seem to interact with the mouse extracellular matrix; Fig. 1). When the behavior of such refined xenograft models was considered against the action of "standard" chemotherapeutic agents, xenografts correctly predicted response in 90% (19 of 21

tumors) and resistance in 97% of the patients (57 of 59 tumors; ref. 14; Fig. 1).

Nonetheless, there are clear limitations to xenograft models. Certain compounds, such as camptothecins, look very good in mice, but owing to class-specific differences in protein binding or metabolism have features where dose escalation in humans is hampered by intrinsic differences between mouse and human toxicity features (16). Given that humans are not mice, the incorporation of pharmacologic and pharmacodynamic end points into the early clinical trials in humans should be considered imperative and will allow an estimate of how likely the human efficacy experience (i.e., phase II) can mirror the mouse efficacy experience once human clinical data become available. In essence, a second major "crucial" point for incorporating mouse xenograft information is not only at the decision making to initially develop a drug for phase I trials but also at the transition from phase I to phase II studies.

Other mouse models, although not proven, could certainly aid in decision making. Transgenic animals, in particular, allow the definition of biological steps in the evolution of a neoplasm. However, in a limited number of available human-like histologies, the penetrance and onset of tumor evolution in such models argues against their use for "screening" purposes (7). One must have already defined the compound of choice with reasonable certainty to optimally use such models and be clear about its pharmacologic features, usually with the prospect of continued nonparenteral modes of administration. A second issue is that the choice of the transgenic model must reliably mirror the target in the human disease to avoid an overestimation of the value of the transgenic result.

Recent efforts have sought to avoid full-scale xenograft experiments for compounds that have little likelihood of performing well from a pharmacologic perspective. These have used human tumor cells grown in "hollow fibers," constructed from materials permeable to most molecular weights of relevant cancer drug candidates. These fibers can be placed in various mouse body compartments and allow several cell types to be studied in one mouse experiment (6, 17). This procedure effectively allows "triage" of compounds. A consistent finding is that more potent compounds *in vitro* have the highest likelihood of activity, an experience also mirrored in the NCI Canada retrospective (12, 13). Although certain limitations of the hollow fiber system have been described, the approach allows the choice to proceed to a xenograft experience to be made in a more informed and economically more considered fashion (18).

Implications

Our collective position is that human tumor xenograft models can be quite useful even in the absence of pharmacologic information. However, there are a number of conditions that should be applied to their use apart from whether the tumors have been directly established from patient samples. First, one needs to have knowledge of the biological features of the model. Not all tumor systems are useful for study, with for example, the

occurrence of "no takes," rendering the statistical definition of activity problematic (6). Guidance and advice from a biostatistical consultant is key and plays to the value of mouse xenografts in contrast to transgenic systems in that appropriately staged tumors can be reliably and economically generated to allow evaluation of a series of compounds (19). Second, as we evolve toward the more routine evaluation of targeted therapeutics, for a new molecular entity whose target is defined, understanding the *in vitro* area under the pharmacologic plasma concentration-time curve producing the desired target effect is an essential piece of information that is not as frequently addressed as it should. This information should be generated together with information about the pharmacology of the compound given by different routes and schedules to non-tumor-bearing mice before undertaking actual antitumor efficacy experiments. In the absence of a pharmacology assay, defining a target effect in the tumor compartment could substitute and indeed serve as the basis or a pharmacodynamic assay for the drug's effect in clinical trials (20). Third, understanding the pharmacodynamic action of the drug in the tumor and a mouse surrogate compartment (e.g., peripheral blood mononuclear cells) in relation to the occurrence of an antitumor effect is of great value, particularly in relation to murine pharmacology. It is the magnitude of that effect in relation to host toxicity that is the ultimate basis for building enthusiasm to move to the clinic. An experience that captures this sequence as it actually occurred was in the development of the proteasome inhibitor bortezomib (21).

In summary, mouse xenograft models should not be viewed as ideal models for cancer drug development. Altered, nonhuman host stroma, poor predictive value when applied in an empirical sense, and questionable relation to the naturally occurring human disease are but a few features, which temper enthusiasm for their use. However, greater value of these models accrues when they are understood well in relation to the pharmacologic and pharmacodynamic attributes of the agent under consideration particularly after human phase I pharmacology information is available. Mouse xenograft models can serve as a useful "filter" for defining the ability of an agent to pass physiologic barriers (e.g., "action at a distance" pharmacologically can be confirmed with oral, i.v., or i.p. administration affecting s.c. tumors), allow selection of a development candidate from an array of congeners, provide a basis for schedule selection in a clinical trial, and allow the design of a clinical trial that is an exact mirror of an experience causing a positive therapeutic effect. Human tumor xenograft data that show activity of an experimental agent are also a positive feature in asking a patient to commit their time and effort to enter a phase I trial. Finally, it is important to note that no clinically approved agent today for the treatment of cancer including targeted agents has lacked activity in conventional preclinical *in vivo* models. We are unaware of any molecule clinically approved as safe and effective, which is devoid of activity in every mouse model tested.

Acknowledgments

Received 10/7/2005; accepted 12/22/2005.

References

1. Gura T. Cancer Models: systems for identifying new drugs are often faulty. *Science* 1997;278:1041-2.
2. Leaf C. The war on cancer: the models of cancer stink. *Fortune* 2004;149:76-88.
3. Petrilli R, Petrilli V, Groff AH, et al. Antifungal efficacy of caspofungin (MK-0991) in experimental pulmonary aspergillosis in persistently neutropenic rabbits: pharmacokinetics, drug disposition, and relationship to galactomannan antigenemia. *Antimicrob Agents Chemother* 2002;46:12-25.

4. Meuwissen R, Linn SC, Linnola M, Zevenhoven J, Mooi JW, Berns A. Induction of small cell lung cancer by somatic inactivation of both Trp53 and Rb1 in a conditional mouse model. *Cancer Cell* 2003;4:181-9.
5. Suggit M, Bibby MC. 30 years of preclinical anticancer drug screening: empirical to target-driven approaches. *Clin Cancer Res* 2005;11:971-81.
6. Alley MC, Hollingshead MG, Dykes DJ, Waud WR. Human tumor xenograft models in NCI drug development. In: Teicher BA, Andrews PA, editors. *Anticancer drug development*. San Diego (CA): Academic Press; 2004. p. 125-52.
7. Burger AM, Fiebig HH. Screening using animal systems. In: Baguley BC, Kerr DJ, editors. *Anticancer drug development*. San Diego (CA): Academic Press; 2001. p. 285-97.
8. Ishii S, Nagasawa M, Kariya Y, Yamamoto H. Selective cytotoxicity of brefeldin A against human tumor cell lines. *J Antibiotics* 1989;42:1877-8.
9. Volpe DA, Tomaszewski JE, Parchment R, et al. Myelotoxic effects of the bifunctional alkylating agent bizelesin on human, canine, and murine myeloid progenitor cells. *Cancer Chemother Pharmacol* 1996; 39:143-9.
10. Scott AM, Wiseman G, Welt S, et al. A Phase I dose-escalation study of sibrizumab in patients with advanced or metastatic fibroblast activation protein-positive cancer. *Clin Cancer Res* 2003;9: 1639-47.
11. Fiebig HH, Burger AM. Human tumor xenografts and explants. In: Teicher BA, editor. *Animal models in cancer research*. Totowa (NJ): Humana Press, Inc; 2001. p. 113-37.
12. Johnson JL, Decker S, Zaharevitz D, et al. Relationships between drug activity in NCI preclinical *in vitro* and *in vivo* models in early clinical trials. *Br J Cancer* 2001;84:1424-31.
13. Voskoglou-Nomikos T, Pater JL, Seymour L. Clinical predictive value of the *in vitro* cell line, human xenograft, and mouse allograft preclinical cancer models. *Clin Cancer Res* 2003;9:4227-39.
14. Fiebig HH, Maier A, Burger AM. Clonogenic assay with established human tumor xenografts: correlation of *in vitro* to *in vivo* activity as a basis for anticancer drug discovery. *Eur J Cancer* 2004;40:802-20.
15. Peterson JK, Houghton PJ. Integrating pharmacology and *in vivo* cancer models in preclinical and clinical drug development. *Eur J Cancer* 2004;40:837-44.
16. Mi Z, Burke TG. Marked interspecies variations concerning the interactions of camptothecin with serum albumins: a frequency domain fluorescence spectroscopic study. *Biochemistry* 1994;33:12540-5.
17. Decker S, Hollingshead M, Bonomi CA, Carter JP, Sausville EA. The hollow fibre model in cancer drug screening: the NCI experience. *Eur J Cancer* 2004;40: 821-6.
18. Shnyder SD, Hassan J, Cooper PA, et al. Development of a modified hollow fibre assay for studying agents targeting the tumor neovasculature. *Anticancer Res* 2005;25:1889-94.
19. Tan M, Fang HB, Tian GL, Houghton PJ. Repeated measures models with constrained parameters for incomplete data in tumour xenograft experiments. *Stat Med* 2005;24:109-19.
20. EORTC Laboratory Research Division [no authors listed]. European Organization for Research and Treatment of cancer (EORTC) Laboratory Research Division workshop on the role of *in vivo* preclinical models in the development of contemporary cancer therapeutics. Verona, Italy: 3 February 2001. *Eur J Cancer* 2001;37: 1976-80.
21. Adams J, Palombella VJ, Sausville EA, et al. Proteasome inhibitors: a novel class of potent and effective antitumor agents. *Cancer Res* 1999;59:2615-22.

Response

The review by Burger and Sausville makes the argument for the use of xenografts in preclinical trials of anticancer therapeutics. Although the nature of the review was to take one position and defend it, it is clear that the authors for both positions are keenly aware of the strengths and weaknesses of both genetically engineered mouse models (GEMs) and xenografts. It is likely that there will be a place for both systems as we move forward in this field. However, there are some specific issues brought up by the authors of the pro xenograft position that deserve comment:

First, the authors state that a weakness of GEMs of cancer is that only a subset of the models shows high degree of histologic similarity to their human counterparts. This may be true; however, those that do show much more similarities than any of the known xenografts. In fact, some GEMs are so similar to human tumors that clinical pathologists have a difficult time distinguishing mouse from human under the microscope. We are not aware of any xenograft models that can be described this way.

Second, the authors assume that the appropriate use of GEMs is to test molecules that target the components that were used to cause the disease. It is certainly true that such studies have been done as proof of principle, but these studies are admittedly a bit of a house of cards and should not be interpreted as directly reflecting the expected response in human tumors. It seems to us that, ultimately, the best use of these models is to test individual or combinations of therapies not central to the oncogenic mechanism of the particular model. This would hopefully place the target of the therapy more accurately in molecular context.

Finally, the authors point out that one critical use of *in vivo* tumor models is the determination of time-dependent plasma concentration versus molecular efficacy within the tumor. We absolutely agree. However, if the histology of the tumor, or its interaction with the organ in which it resides does not reflect the human counterpart the data could be misleading. In this regard, the histologically accurate GEMs are far more likely to provide such information in a way that reflects the human situation. An extreme example of this would be a diffuse glioma arising within the brain behind the blood-brain barrier compared with a xenograft implanted into the flank or even orthotopically implanted into the brain.

In spite of the long history and ease of use, concerns remain about tumor xenograft assays with regard to gaining mechanistic insight, appropriate immune context or tissue setting, and the use of tumor cell lines with an uncertain molecular relationship to the original human tumors from which they were derived. Further research and time will tell us in what specific circumstances xenografts or genetically engineered mice are the better choices for preclinical drug trials, and how the data for each should be interpreted.

Oren J. Becher

Eric C. Holland

Departments of Cancer Biology and Genetics, Surgery (Neurosurgery) and Neurology, and Pediatrics Memorial Sloan-Kettering Cancer Center New York, New York